

Abstract

A method, computer program and process for administering the workload of a database system as it executes one or more requests are disclosed. The one or more requests are sorted into one or more workload groups. Each workload group has an associated level of service desired from the database system. The one or more requests are executed in an order intended to achieve the levels of service associated with each of the workload groups. The system resources are assigned to the one or more workload groups as necessary to provide the level of service associated with each workload group. The execution of requests is monitored on a short-term basis to detect a deviation from the level of service greater than an short-term threshold. If such a deviation is detected, the assignment of system resources to workload groups is adjusted to reduce the deviation. Monitoring is also performed on a long-term basis to detect deviations from the expected level of service greater than a long-term threshold. If such a deviation is detected, the execution of requests is adjusted to better provide the expected level of service.